Conformity enhances network reciprocity in evolutionary social dilemmas

Attila Szolnoki¹ and Matjaž Perc²,³,⁴

¹Institute of Technical Physics and Materials Science, Research Centre for Natural Sciences, Hungarian Academy of Sciences, PO Box 49, 1525 Budapest, Hungary
²Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia
³Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia
⁴CAMTP-Center for Applied Mathematics and Theoretical Physics, University of Maribor, Kranova 2, 2000 Maribor, Slovenia

The pursuit of highest payoffs in evolutionary social dilemmas is risky and sometimes inferior to conformity. Choosing the most common strategy within the interaction range is safer because it ensures that the payoff of an individual will not be much lower than average. Herding instincts and crowd behaviour in humans and social animals also compel to conformity in their own right. Motivated by these facts, we here study the impact of conformity on the evolution of cooperation in social dilemmas. We show that an appropriate fraction of conformists within the population introduces an effective surface tension around cooperative clusters and ensures smooth interfaces between different strategy domains. Payoff-driven players brake the symmetry in favour of cooperation and enable an expansion of clusters past the boundaries imposed by traditional network reciprocity. This mechanism works even under the most testing conditions, and it is robust against variations of the interaction network as long as degree-normalized payoffs are applied. Conformity may thus be beneficial for the resolution of social dilemmas.

1. Introduction

Not only are social interactions limited and thus best described not by well-mixed models but rather by models entailing networks [1–5], it is also a fact that these interactions are not always driven by a selfish agenda such as fitness maximization [6,7]. Although high individual fitness, which is most often quantified by a scalar payoff value, is at the heart of success in evolutionary games [8–13], social interactions are also often aimed at fostering the general sense of belonging and the identification with a particular group or a way of thinking, with a subculture or even with a fashionable trend or movement [14]. And one of the most common actions for achieving this is simply to comply or to conform with the most widespread and established ideas or ideals of the reference group to which one aspires to.

Despite the fact that payoff maximization is a surprisingly apt description of interactions among simpler forms of life, such as bacteria and plants [15,16] as well as among viruses [17], and thus rightfully permeates evolutionary game theory, the consideration of alternative targets, especially for interactions among humans and social animals, appears to be justified. Of particular relevance in this case are social dilemmas, where the interests of individuals are at odds with what is best for the society as a whole, and none has received as much attention as the Prisoner’s Dilemma game [18–37]. Each instance of the game is contested by two players who have to decide simultaneously whether they want to cooperate or defect. The dilemma is given by the fact that although mutual cooperation yields the highest collective payoff, a defector will do better if the opponent decides to cooperate. A purely rational payoff-driven player should thus always decide to defect.

Ample research has already been devoted to the identification of mechanisms that may lead to a cooperative resolution of social dilemmas. Classic mechanisms
are reviewed in reference [38], among which network reciprocity owing to Nowak & May [39] has motivated an impressive array of studies aimed at understanding the evolution of cooperation in structured populations [40–44]. Methods of statistical physics have proven particularly suitable for this task, as evidenced by the seminal studies of the evolution of cooperation on small-world [45,46], scale-free [47,48], coevolving [49,50], hierarchical [51], bipartite [52] and most recently also on multilayer networks [53–57]. Moreover, many coevolutionary rules [42] have been introduced that may generate favourable interaction networks spontaneously [29,58–64]. There is also experimental evidence in favour of the fact that a limited interaction range does play a prominent role by the evolution of cooperation [65,66], especially so if coupled with rewiring [67].

The key assumption behind existing research, however, has been that every player aspires only to maximizing its own payoff, although this is obviously not always the case. In fact, there exists compelling evidence in favour of the fact that conformity also plays an important role [14], especially among humans and social animals. The consideration of conformity dictates the adoption of the strategy that is most common within the interaction range of the player, regardless of the expected payoff [68,69]. By adopting the most common strategy, the conformists thus coordinate their behaviour in a way that minimizes individual risk and fosters coherence within the population. Unlike previous research, however, we do not assume that this ‘cultural transmission’ affects all the players in the population uniformly [69]. Instead, we take into account the fact that players are diverse in their aspirations, and that thus some are keen on maximizing their payoffs, whereas others are simply content to adopt the most common strategy in their neighbourhood. Interestingly, such behaviour has recently also been observed in the realm of an economic experiment involving the public goods game with institutionalized incentives [70].

The question that we therefore wish to address in the continuation is: what is the impact of conformity-driven players on the evolution of cooperation in evolutionary social dilemmas? In particular also, what role does the fraction of conformity-driven players within a population play? One may expect that conformity-driven players will push the system towards neutral evolution, especially when they represent the majority of the population. But interestingly, this is not always the case. In what follows, we show that the introduction of conformists to the Prisoner’s Dilemma game enhances network reciprocity, and thus favours the evolution of cooperation. In particular, we demonstrate how conformity-driven players introduce spontaneous flocking of cooperators into compact clusters with smooth interfaces separating them from defectors. Furthermore, we elaborate on the responsible microscopic mechanisms, and we also test the robustness of our observations. Taken together, we provide firm evidence in support of conformity-enhanced network reciprocity and show how conformists may be beneficial for the resolution of social dilemmas. First, however, we proceed with presenting the details of the mathematical model.

2. Evolutionary games with conformists

We study evolutionary social dilemmas on the square lattice and the Barabási–Albert scale-free network, each with an average degree $k = 4$ and size $N$. For the generation of the scale-free network, we implement the standard growth and preferential attachment algorithm [71]. Accordingly, starting from a small number of vertices ($m_0 = 3$), a new node with $m = 2$ edges is connected to an existing node $x$ with probability $P(k_x) = k_x / \sum y k_y$, where $k_x$ denotes the degree of node $x$. This growth and preferential attachment scheme yields a network with an average degree $k_w = 2m$, and a power-law degree distribution with the slope of the line equaling $\approx -3$ on a double-logarithmic scale. The two considered networks are representative of the simplest homogeneous and the strongly heterogeneous interaction topology.

Each player is initially designated either as cooperator (C) or defector (D) with equal probability, and each instance of the game involves a pairwise interaction where mutual cooperation yields the reward $R$, mutual defection leads to punishment $P$, and the mixed choice gives the cooperators the sucker’s payoff $S$ and the defector the temptation $T$. We predominantly consider the weak Prisoner’s Dilemma, such that $T > 1$, $R = 1$ and $P = S = 0$, but we also consider the true Prisoner’s Dilemma in the form of the donation game, where $T = 1 + b$, $R = 1$, $P = 0$ and $S = -b$.

We simulate the evolutionary process in accordance with the standard Monte Carlo simulation procedure comprising the following elementary steps. First, according to the random sequential update protocol, a randomly selected player $x$ acquires its payoff $P_x$ by playing the game with all its neighbours. Next, player $x$ randomly chooses one neighbour $y$, who then also acquires its payoff $P_y$ in the same way as previously player $x$. Once both players acquire their payoffs, then player $x$ adopts the strategy $s_x$ from player $y$ with a probability determined by the Fermi function

$$\Gamma(P_x - P_y) = \frac{1}{1 + \exp((P_x - P_y)/K)}, \quad (2.1)$$

where $K = 0.1$ quantifies the uncertainty related to the strategy adoption process [40,72]. In agreement with previous works, the selected value ensures that strategies of better-performing players are readily adopted by their neighbours, although adopting the strategy of a player that performs worse is also possible [73,74]. This accounts for imperfect information, errors in the evaluation of the opponent, and similar unpredictable factors.

To introduce conformity, we designate a fraction $\rho$ of the population as being conformity-driven, and this influences the strategy adoption rule. In particular, each conformist $x$ simply prefers to adopt the strategy that is most common within its interaction range. Equation (2.1) thus no longer applies. Instead, if player $x$ is a conformist, we use

$$\Gamma(N_s - k_x) = \frac{1}{1 + \exp((N_s - k_x)/K)}, \quad (2.2)$$

where $N_s$ is the number of players adopting strategy $s_x$ within the interaction range of player $x$, whereas $k_x$ is one half of the degree of player $x$. It is worth pointing out that the application of equation (2.2) results in the conformity-driven player adopting, with a very high probability, whichever strategy (either C or D) is at the time the most common in its neighbourhood. Nevertheless, it is still possible, yet very unlikely that a conformist will adopt the strategy that is in the minority. If, however, the number of cooperators and defectors in the neighbourhood is equal, the conformity-driven player will change its strategy with probability 1/2.

In terms of the simulation procedure, we note that each full Monte Carlo step (MCS) consists of $N$ elementary steps
described above, which are repeated consecutively, thus giving a chance to every player to change its strategy once on average. All simulation results are obtained on networks typically comprising \( N = 10^5 \) to \( 10^6 \) players, although the usage of larger networks is necessary in the proximity to phase transition points. We determine the fraction of cooperators \( f_C \) in the stationary state after a sufficiently long relaxation time lasting up to \( 10^5 \) MCS. To further improve accuracy, the final results are averaged over 400 independent realizations, including the generation of the scale-free networks and random initial strategy distributions, for each set of parameter values.

### 3. Results

Before presenting the results in structured populations, we summarize briefly the results in well-mixed populations. In the Prisoner’s Dilemma game, defectors dominate completely in the absence of conformity-driven players. But if all players are conformists, then everybody loses interest in payoffs, and the evolution is simply a random drift. Consequently, the system may terminate into an all-C or an all-D phase, ultimately yielding an average fraction of cooperators \( f_C = 0.5 \). The intermediate region is more interesting where two different cases have to be considered separately. If \( \rho < 0.5 \), the majority of players is still driven by payoffs, and the system thus eventually terminates into an all-D phase. If \( \rho > 0.5 \), however, the majority of players is conformity-driven, meaning that either an all-C or an all-D phase will ultimately emerge among them, thus giving rise to \( f_C = 0.5 \cdot \rho \). In sum, in the low \( \rho \) region, defectors always dominate, whereas in the high \( \rho \) region, the average cooperation level can never exceed 0.5, and this is regardless of the value of \( T \).

In structured populations, the presence of conformity-driven players has significantly more unexpected and even counterintuitive consequences. We begin by presenting results obtained with the weak Prisoner’s Dilemma on the square lattice. The colour map presented in figure 1 encodes the stationary fraction of cooperators \( f_C \) in dependence on the temptation to defect \( T \) and the density of defectors \( \rho \). The colour map encodes the stationary fraction of cooperators in dependence on the temptation to defect \( T \) and the density of defectors \( \rho \). The role of players seeking to maximize their payoffs is to reveal the expansion of clusters in the socially desirable direction.

To illustrate the microscopic dynamics behind the described cooperation, we show in figure 2 a series of characteristic strategy distributions that describe the time evolution of the game from a random initial state. For the sake of clarity, we use different colours not only for cooperators (blue) and defectors (red), but also for distinguishing conformity-driven (pale) and payoff-driven (dark) players. From the outset, defectors spread very efficiently, and indeed, only a small flock of cooperators is able to survive (figure 2b). The compact cluster protects the cooperators from extinction, and this is in fact a very typical time evolution for an evolutionary game that is contested in a structured population. What distinguishes this cluster from an ordinary cluster that would be due solely to network reciprocity is its smooth interface that separates the competing domains. In fact, the dynamics between conformity-driven players is conceptually similar to the so-called majority-voter model [75]. It is easy to see that the larger the fraction of conformity-driven players, the smoother the interface between the competing domains. Put differently, the propensity of players to comply introduces an effective surface tension around cooperative clusters that is completely absent in traditional voter models, where rough interfaces and slow coarsening are common [76].

![Figure 1. Evolution of cooperation in the weak Prisoner’s Dilemma with conformity-driven players, as obtained on the square lattice in dependence on the temptation to defect \( T \) and the density of defectors \( \rho \). The colour map encodes the stationary fraction of cooperators \( f_C \) in dependence on \( T \) and \( \rho \).](http://rsif.royalsocietypublishing.org/)

When payoff-driven players are rare, the symmetry-breaking owing to them can be analysed in more detail by

| \( T \) \( \rho \) \( f_C \) |
|---|---|---|
| 0.2 | 0.2 | 0.2 |
| 0.4 | 0.4 | 0.4 |
| 0.6 | 0.6 | 0.6 |
| 0.8 | 0.8 | 0.8 |
| 1.0 | 1.0 | 1.0 |

\( f_C \) increases above zero, the neutral strategy evolution at \( \rho = 0 \) gives rise to a bell-shaped, non-monotonic dependence of \( f_C \) on \( \rho \), which is particularly pronounced at intermediate values of \( T \). Based on the results presented in figure 1, we may thus conclude that it is beneficial for the whole society if the majority of the population consists of conformity-driven players. Nevertheless, a certain fraction of payoff-driven players is necessary to induce symmetry-breaking along the interfaces that separate competing strategy domains. The role of conformists is hence simply to homogenize the population locally, whereas the role of players seeking to maximize their payoffs is to reveal the long-term benefits of cooperation and thereby to guide the expansion of clusters in the socially desirable direction.

When payoff-driven players are rare, the symmetry-breaking owing to them can be analysed in more detail by
Zooming in on the most likely elementary steps along the interface that separates cooperators and defectors, as shown in the inset of figure 3. In the absence of payoff-driven players (dark colours), the most stable interface separating conformity-driven players (pale colours) is a straight line where the most likely change happens at a step. The position of this step propagates randomly until a payoff-driven player is encountered, at which point imitation becomes a possibility. The most probable elementary steps are depicted in the inset of figure 3. As noted, conformity-driven players at the step, marked by tilted lines, are able to change their strategy with probability 1/2. More interestingly, payoff-driven players are able to adopt another strategy, yet they cannot pass their strategy by using their ‘success’, because neighbouring conformity-driven players do not care about higher payoffs. Accordingly, the most probable invasions are only those marked by white arrows. By summing up these elementary processes, we can estimate the time variation of the fraction of cooperators according to

$$\frac{\Delta f_C}{\Delta t} = \left( \Gamma(2T - 3R) - \frac{1}{2} \right) + \left( \frac{1}{2} - \Gamma(2R - 2T) \right).$$

(2.3)

where the $\Gamma$ function is defined as in equation (2.1). We plot $\Delta f_C/\Delta t$ in dependence on the temptation to defect in figure 3.

The result suggests that cooperation will spread below a threshold value, even if $T > R$, as a consequence of the broken symmetry described above. Indeed, this simple approximation is able to explain why there is a relatively sharp transition between the full C and the full D evolutionary outcome in the high $\rho$ region (figure 1).

We proceed by testing the robustness of conformity-enhanced network reciprocity, first by considering an alternative formulation of the social dilemma. So far, to obtain results that are comparable with previous related works [77,78], we have focused on the weak Prisoner’s Dilemma, which does not constitute the most adverse conditions for the successful evolution of cooperation, because the punishment for mutual defection and the suckers payoff are equal ($P = S = 0$). To amend this, we consider the donation game, where the payoff ranking $T > R > P > S$ corresponds to the true Prisoner’s Dilemma. The consideration of this game is all the more interesting, because network reciprocity alone is virtually unable to sustain cooperation under such testing circumstances [39]. As for the weak Prisoner’s Dilemma in figure 1, for the donation game too, we present a colour map that encodes the stationary fraction of cooperators $f_C$ in dependence on the temptation to defect $T$ and the fraction of conformity-driven players $\rho$ in figure 4. It can be

Figure 2. Evolution of cooperation from a random initial state under the influence of conformity. Depicted are characteristic spatial patterns, as obtained with the weak Prisoner’s Dilemma game on a square lattice using $T = 1.45$ and $\rho = 0.81$. Payoff-driven cooperators (defectors) are depicted dark blue (dark red), while conformity-driven cooperators (defectors) are depicted bright blue (pale red). Starting from a random initial state (a), conformity-driven players introduce spontaneous flocking of cooperators into compact clusters with smooth interfaces separating them from defectors (b). Payoff-driven players subsequently reveal the long-term benefits of cooperation and the cluster grows, all the while maintaining surface tension and thus a smooth interface (c). The effectiveness of this conformity-enhanced network reciprocity eventually propels cooperators to near-complete dominance (d). For clarity, we have here used a small square lattice with linear size $L = 100$. (Online version in colour.)
observed that, even in the most challenging social dilemma, conformity-enhanced network reciprocity is able to ensure widespread dominance of cooperators at remarkably high values of $T$, as long as the value of $r$ is sufficiently large. However, as $r \to 1$, the lack of payoff-driven players introduces the status quo among conformity-driven players, and again the average fraction of cooperators drops to $f_C \approx 0.5$ and thus gives rise to the bell-shaped, non-monotonous dependence of $f_C$ on $r$.

Lastly, we explore the robustness of conformity-enhanced network reciprocity to changes in the topology of the interaction network. Because the square lattice is representative of regular, homogeneous interaction networks, the most interesting test involves considering the highly heterogeneous scale-free network. First, we have to emphasize that the consideration of absolute payoffs on strongly heterogeneous networks already provides ample support to network reciprocity, in particular by ensuring homogeneous strategy ‘clouds’ around hubs [22,47]. The introduction of conformity-driven players is therefore either negligible or even negative. Namely, if a defective hub is designated as conformity-driven, then it is almost impossible to revert its strategy to cooperation, because the large number of like-minded defective followers will always assure the hub it already conforms perfectly with the neighbourhood. Because previous research has highlighted the main cooperator-promoting mechanism on scale-free networks is based on the fact that defective hubs eventually become weak and thus vulnerable to strategy change [40], the introduction of conformists can be a notable drawback because this key mechanism becomes disabled.

However, because the application of absolute, cumulative payoffs on strongly heterogeneous interaction networks already raised questions during the early stages of research on this subject [79–81], in particular in the sense that a player might be unable to maintain a large number of connections for free, the application of degree-normalized payoffs was proposed. It was shown that the application of such payoffs erases the ability of heterogeneous networks to sustain large homogeneous cooperative clusters around hubs, and that the cooperation levels return to those observed earlier on regular networks and lattices. With this, we arrive yet again at conditions where the presence of conformity-driven players might help cooperation significantly, and indeed, the results presented in figure 5 fully confirm this expectation. All qualitative features remain the same as by the consideration of the weak Prisoner’s Dilemma and the donation game on the square lattice (compare with figures 1 and 4), and with this we conclude that conformity-enhanced network reciprocity is also robust against variations of the interaction network as long as degree-normalized payoffs are applied.

4. Discussion

We have studied the evolution of cooperation in social dilemmas where a fraction of the population has been designated as being driven by conformity rather than payoff maximization. Unlike in traditional evolutionary game theory, these conformists are no longer concerned with maximizing their payoffs by selecting the most promising strategy for future interactions. Instead, conformity-driven players simply adopt whichever strategy is most common within their interaction range at any given time. We have shown that the presence of conformity-driven players enhances network reciprocity, and thus aids the favourable resolution of social dilemmas. The effectiveness of conformity to do so, however, depends on the fraction of conformists within the population. If the later
are neither too rare nor too common, the flocking of cooperators into compact clusters with smooth interfaces emerges spontaneously. But because the strategy preference among conformists is neutral, a certain fraction of payoff-driven players is necessary to induce symmetry-breaking along the interfaces that separate competing domains. Based on this, we have demonstrated that cooperative clusters are able to expand past the boundaries imposed by traditional network reciprocity, in particular, because the dynamics between conformity-driven players is conceptually similar to the so-called majority-voter model [75]. We have emphasized that the larger the fraction of conformity-driven players, the smoother the interface between the competing domains. In other words, the tendency of conformists to blend in introduces an effective surface tension that does not exist in the traditional voter model where rough interfaces and slow coarsening can be observed [76].

We have shown that the newly identified mechanism is robust to variations of the contested social dilemma, and that it works even under the most testing conditions where traditional network reciprocity completely fails to sustain cooperative behaviour. Moreover, we have shown that conformity promotes cooperation, regardless of the properties of the interaction network, as long as degree-normalized payoffs are applied. If absolute payoffs are applied on strongly heterogeneous networks, then the heterogeneity alone provides the maximal support to network reciprocity, and hence the impact of conformity-driven players is either negligible or even negative. If namely by chance a defective hub is designated as conformity-driven, then it is almost impossible to revert its strategy to cooperation, and this thus disables the key mechanism that ensures elevated levels of cooperation in heterogeneous networks [40].

The presented research has been motivated by the fact that payoff maximization alone is often not the primary goal of social interactions. Unlike interactions among simpler forms of life, interactions among humans and social animals are often driven by a desire to belong or to ‘fit in’ [14]. To conform is thereby simply a frequently used way for achieving this goal, and it is interesting and to a degree counterintuitive to discover that conformity may actually promote the evolution of cooperation. Our results of course take nothing away from payoff maximization as an apt and in fact comprehensive motivator of interactions among bacteria, plants and viruses, but they do suggest that conformity might have had an evolutionary origin in as far as it furthers prosocial behaviour. Furthermore, in addition to possible emotional origins of being a conformist, there might be cases when this preference is actually a payoff-maximizing strategy (by avoiding punishment, for instance). Hence, conformism may be a good heuristic in social decision making [82,83].

We conclude by noting that our model is just an initial step towards the introduction of ‘multi-target’ evolutionary games, which ought to properly take into account the diversity of not only the applied strategies, but also the diversity of individual targets one may hope to achieve by adopting them. An interesting direction for future research might involve players being able to change their motivation over time, for example through ‘cultural transmission’ [68]. Relative timescales in evolutionary dynamics [84] could also play an important role, in the sense that the typical time for ‘motivational change’ might be different from the typical time in which players change their strategy. To explore the consequences of these options appears to be an exciting venture with many relevant implications, and we hope that this paper will motivate further research along this line in the near future.

Funding statement. This research was supported by the Hungarian National Research Fund (grant no. K-101490), the Slovenian Research Agency (grant no. P5-0027) and by the Deanship of Scientific Research, King Abdullah University (grant no. 76-130-35-HiCi).

References


